Training session on Regional Climate Model Evaluation System (RCMES)

Kyo Lee (huikyo.lee@jpl.nasa.gov)
(Jet Propulsion Laboratory, California Institute of Technology)

December 24, 2018

http://rcmes.jpl.nasa.gov

http://climate.apache.org

Acknowledgement

- My special thanks of gratitude to Dr. Sanjay, Dr. Krishnan, Dr. Neena, and Dr. Suhas.
- Regional Climate Model Evaluation System (RCMES) team

  Duane Waliser (PI), Huikyo Lee (co-I), **Alexander Goodman**, Peter Gibson, Elias Massoud, Brian Wilson, Paul Loikith[2], and Antonio Monge[3]

  [1]JPL/Caltech, [2]California State U. LA, [3]Portland State U.
- **Virtual Information-Fabric Infrastructure (VIFI)** team led by Prof. William Tolone at U. of North Carolina, Charlotte
- Parallelized BCSD codes from Dr. TJ Vandal and NASA Earth eXchange (NEX) team at NASA Ames center

RCMES Training Outline (10:00-12:30)

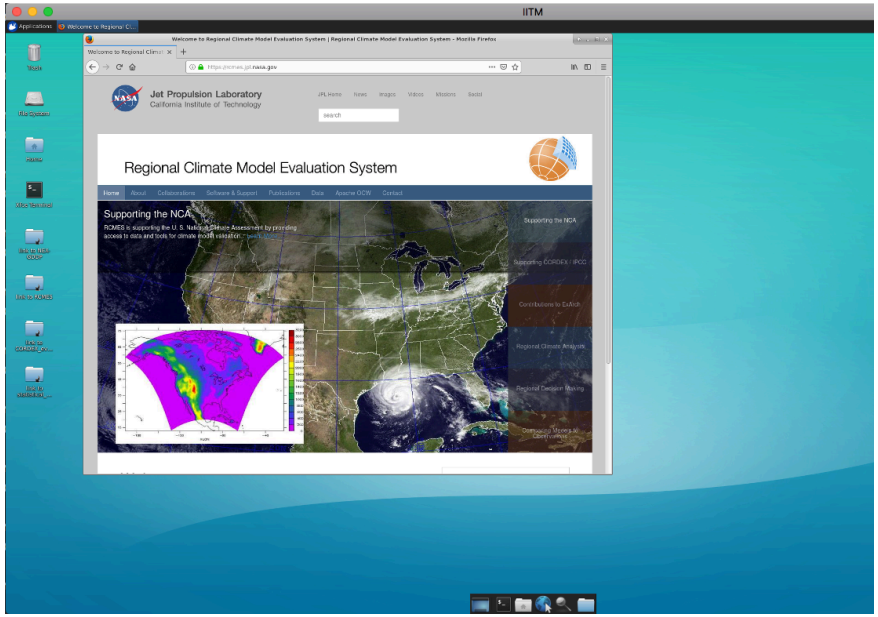
Time	Agenda Item	Process/presentations/materials
10:00-10:15	Welcome and connect to Amazon Elastic Compute Cloud (EC2)	• Check the IP address of the assigned server • Connect to the server using Microsoft Remote Desktop (or terminal software)
10:15-10:30	Activity #1 : Correct biases in CORDEX RCM simulations	• Quantile-based bias correction of the CORDEX WAS simulations using satellite-based precipitation observation data
10:30-10:50	Activity #2 : Evaluate CORDEX RCM simulations	• Systematic evaluation of CORDEX RCMs against obs4mips using RCMES
		Break (10:50-11:00)
11:00-11:30	Activity #3 : Pointwise Statistical downscaling using RCMES	• CMIP5 temperature and precipitation datasets for present and future climate • Compare the IPCC climate change scenarios • (RCP 4.5 vs. RCP 8.5)
11:30-11:45	Activity #4 : Download and visualize the NEX-GDDP data	• NASA Earth Exchange Globally Daily Downscaled Projections (NEX-GDDP) in Amazon Simple Storage Service (S3)
11:45-	Activity #5	• Analyze the bias corrected RCM output and check the evaluation result

Two different ways to connect to the virtual Linux machine on Amazon Web Service

- SSH connection using your terminal application
- Prerequisite software
  - terminal: putty, xshell, xterm
  - X Server: Xming, XQuartz
  - NetCDF/HDF viewer: Panoply
  - (Optional) sftp client: xftp, FileZilla

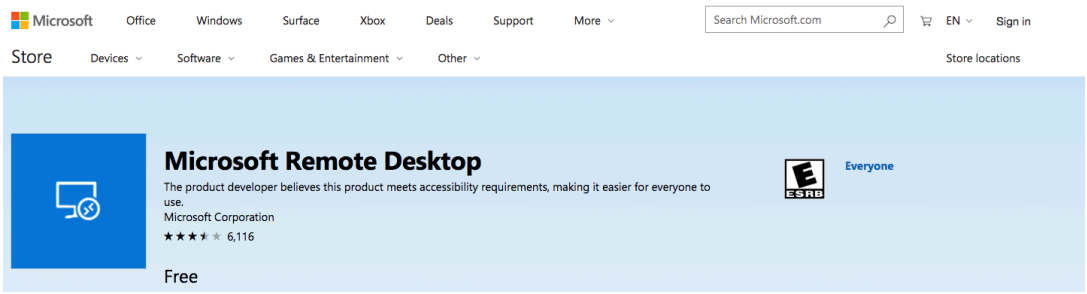
- ssh -Y user1@xx.xxx.xx.xxx
- password: cordex

- Remote desktop



Prerequisite software to run remote desktop

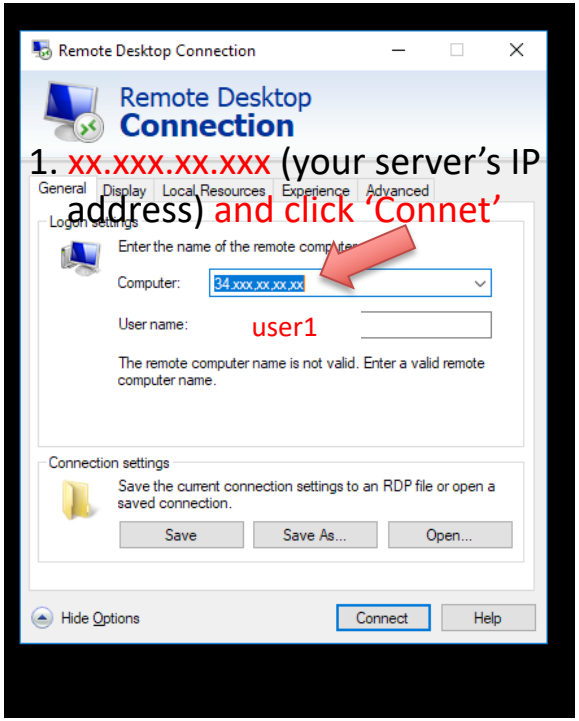
- ~~Linux based system~~
- Windows laptops: Microsoft Remote Desktop



- Macbooks: Microsoft Remote Desktop **10**

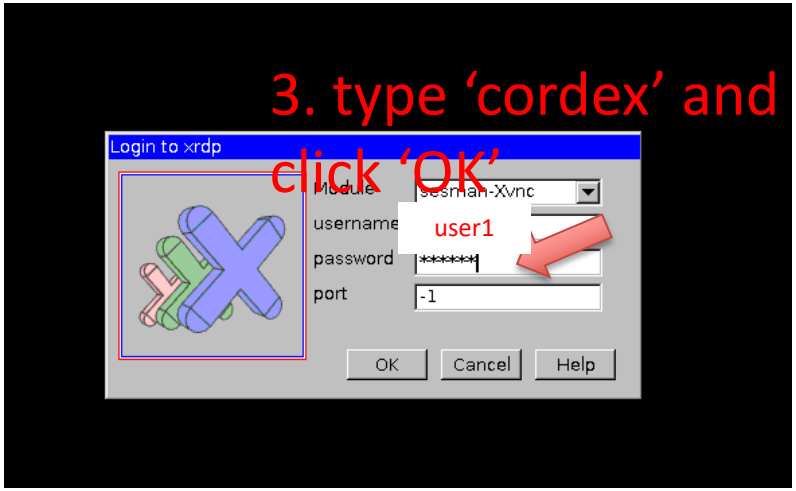
(do not use version 8)

Set up your remote desktop (Windows)



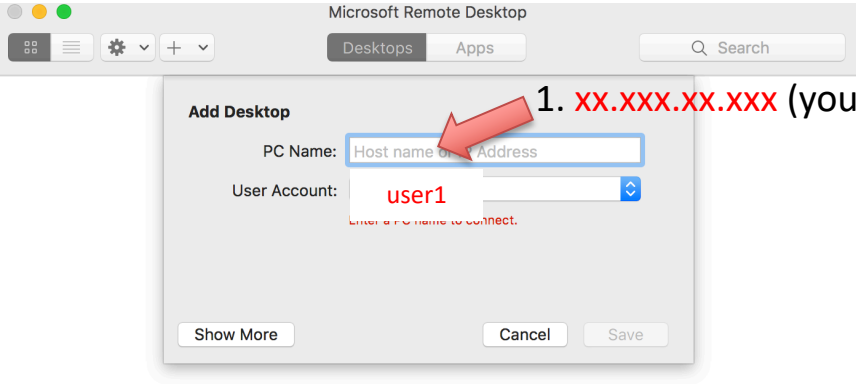
1. xx.xxx.xx.xxx (your server's IP address) and click 'Connet'

2. Click 'OK'

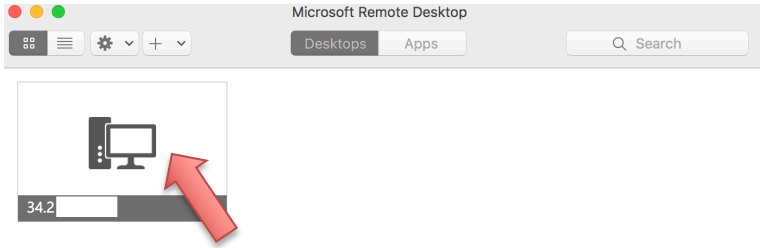


3. type 'cordex' and click 'OK'

Set up your remote desktop (Mac)

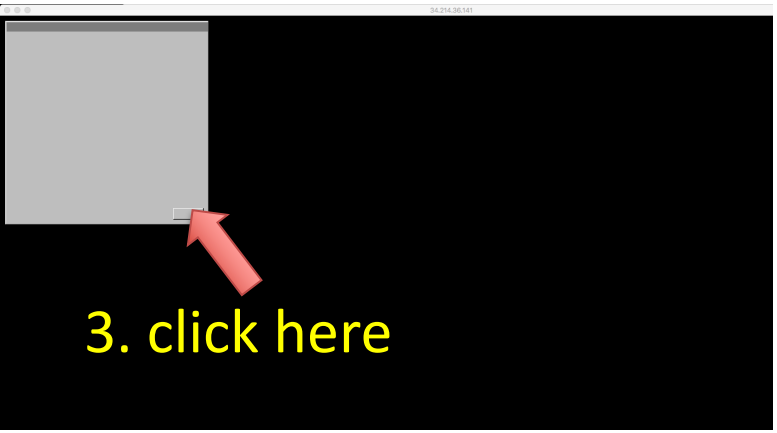


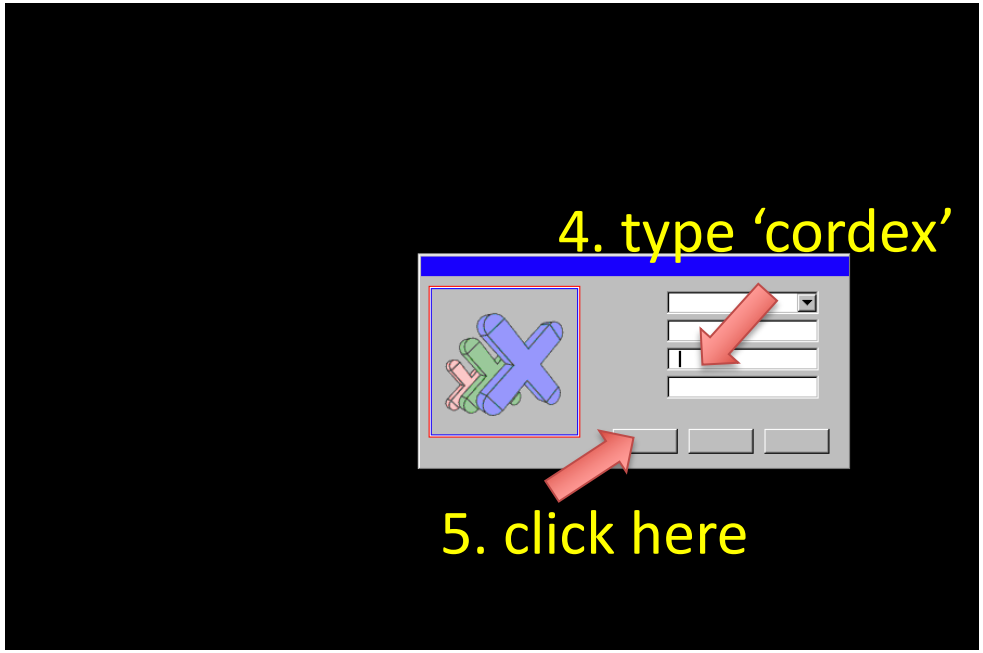
1. xx.xxx.xx.xxx (your server's IP address)



2. Double click!!!

Bugs in the Mac version





- **Activity #1**

: Correct biases in CORDEX RCM simulations

- Activity #2

: Evaluate CORDEX RCM simulations

- Activity #3

: Pointwise Statistical downscaling using RCMES

- Activity #4

: Download and visualize the NEX-GDDP data

Running the bias correction script

(courtesy of Dr. TJ Vandal at NASA Ames, https://github.com/tjvandal/bcsd-python)

1. Open Terminal and type `cd RCMES`

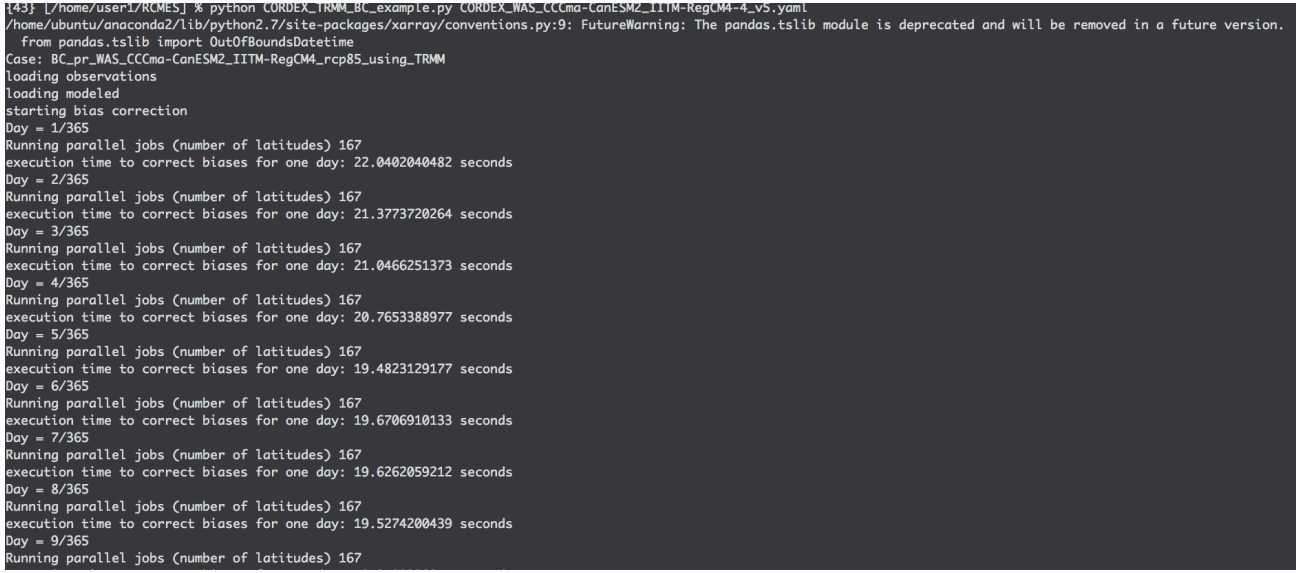
(Five options: please choose one of yaml files)

ex) To correct biases in the CSIRO-MK3-6-0_IITM-RegCM4-4_v5 for the CORDEX South Asia domain,

Python script — One of the five configuration files

```
python CORDEX_TRMM_BC_example.py CORDEX_WAS_CSIRO-QCCCE-CSIRO-Mk3-6-0_IITM-RegCM4-4_v5.yaml
```

(Running this parallelized script takes more than an hour.)



Bias Correction of CORDEX simulations

- The IITM-RegCM simulations have high spatial resolution (~44 km) relative to CMIP GCMs.

- ~~BCSD~~ => BC : spatial disaggregation (SD) may not be necessary thanks to the resolution of CORDEX simulations.

Quantile mapping to correct simulated precipitation using TRMM observations (1)

- Inside the configuration file (CORDEX_WAS_CCCma-CanESM2_IITM-RegCM4-4_v5.yaml)

fobserved: TRMM_regridded_RegCM4-4_v5_day_19980101-20131201_WAS-44.nc

observed_varname: TRMM_daily_pr

fmodeled_present: pr_WAS-44_CCCma-CanESM2_historical_r1i1p1_IITM-RegCM4-4_v5_day_19900101-20051231.nc

fmodeled_future: pr_WAS-44_CCCma-CanESM2_rcp85_r1i1p1_IITM-RegCM4-4_v5_day_20840101-20991231.nc

modeled_varname: pr

(Observation)
Read TRMM_daily_pr from TRMM_regridded_RegCM4-4_v5_day_19980101-20131201_WAS-44.nc

(Simulation for the present climate)
Read pr from pr_WAS-44_***_19900101-20051231.nc

(Simulation for the future climate)
Read pr from pr_WAS-44_***_20840101-20991231.nc

Quantile mapping to correct simulated precipitation using TRMM observations (2)

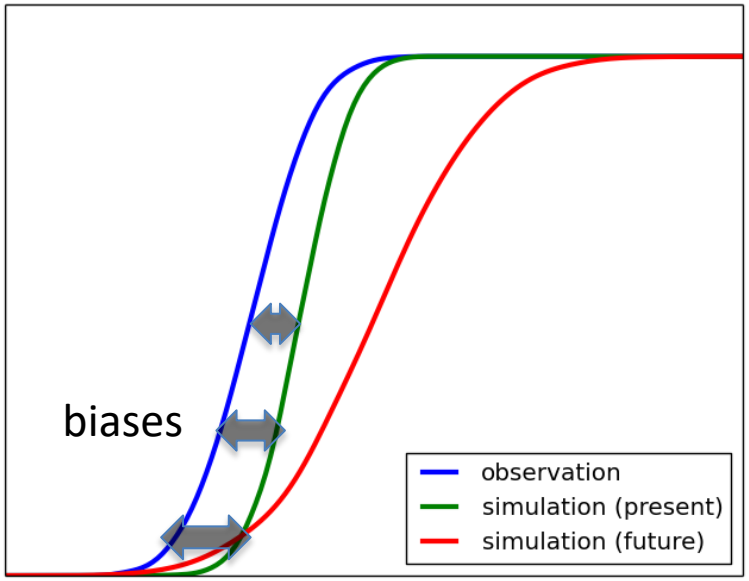
- At each RCM grid point, biases in simulated precipitation are is corrected for each quantile (0.5-99.5%) by comparing two cumulative distributions from TRMM and the RCM (±15 days).

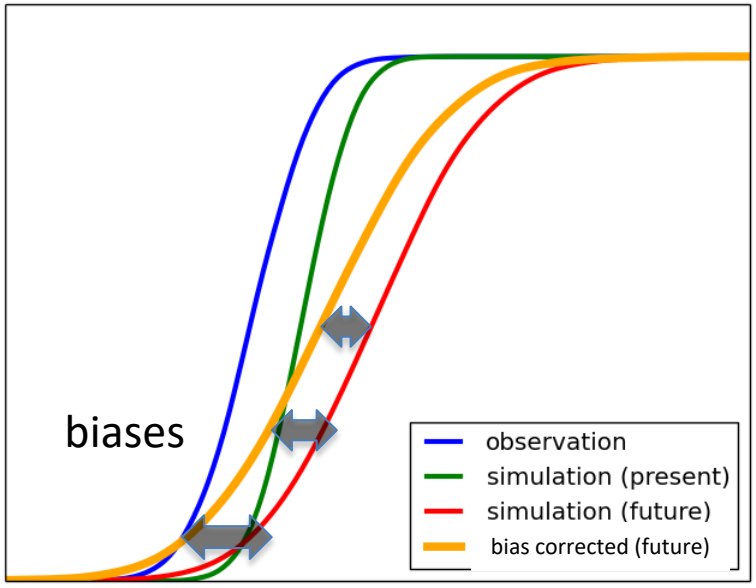
(Observation)
TRMM_daily_pr for 19980101-20131231

(Simulation for the future climate)
pr_WAS-44_***_20840101-20991231.nc

(Simulation for the present climate)
pr_WAS-44_***_19900101-20051231.nc

(Bias corrected future simulation)
BC_pr_WAS-44_***_20840101-20991231.nc





- Activity #1

: Correct biases in CORDEX RCM simulations

- **Activity #2**

: Evaluate CORDEX RCM simulations

- Activity #3

: Pointwise Statistical downscaling using RCMES

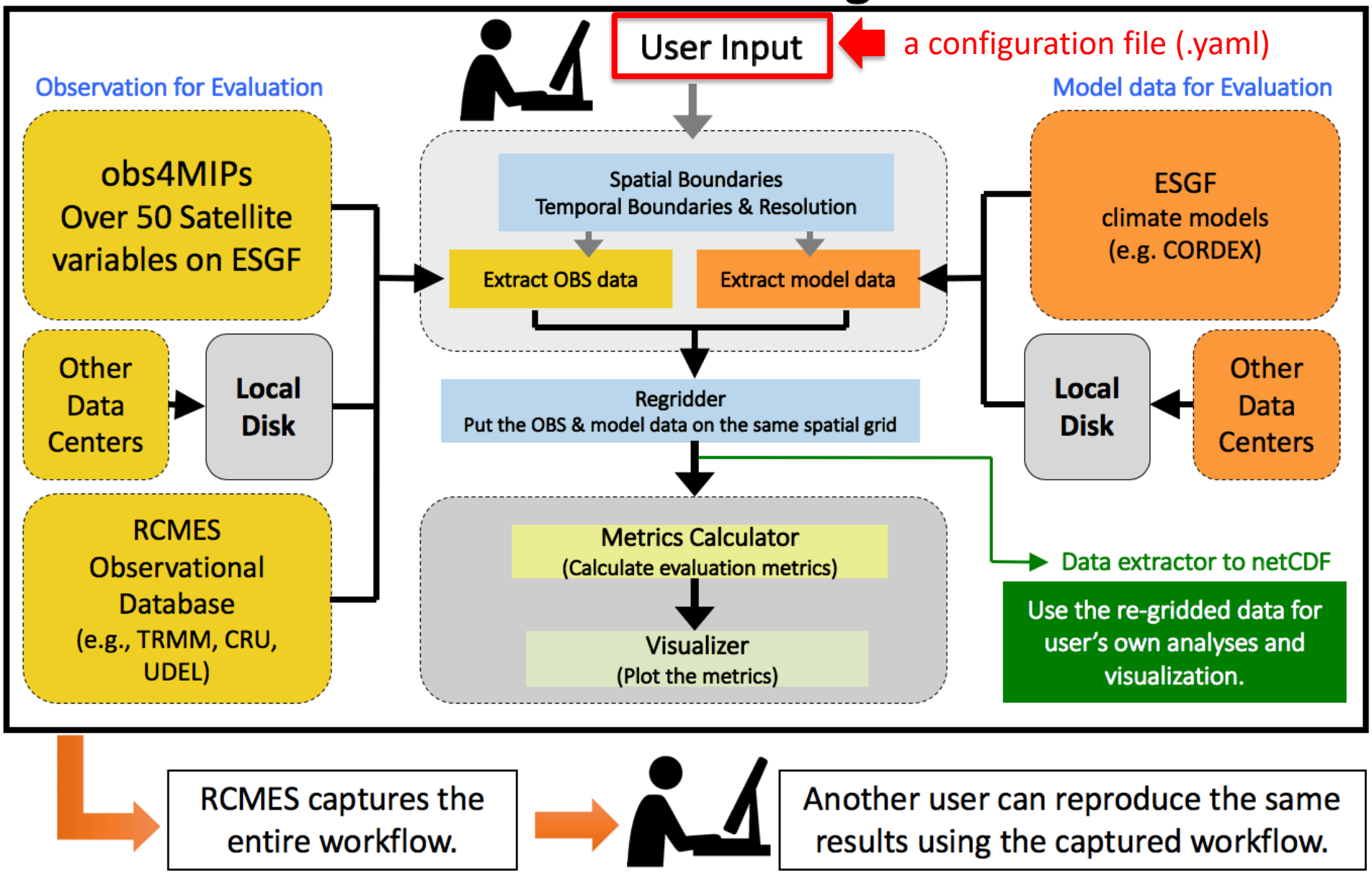
- Activity #4

: Download and visualize the NEX-GDDP data

The Regional Climate Model Evaluation System (RCMES, https://rcmes.jpl.nasa.gov)

- Lee et al. (2018), Regional Climate Model Evaluation System powered by Apache Open Climate Workbench v1.3.0: an enabling tool for facilitating regional climate studies, Geoscientific Model Development.
- Python-based open source software powered by the Apache Open Climate Workbench (OCW)
- Main components
  1) Database of observations
  2) Toolkit for facilitating systematic evaluation of CORDEX RCMs using satellite observations (Activity #2)
  3) Statistical downscaling of coarse-resolution GCM output or bias correction of high-resolution RCM output (Activity #1 & 3)
  4) Stand-alone scripts for data processing and visualization based on OCW

Model evaluation using RCMES

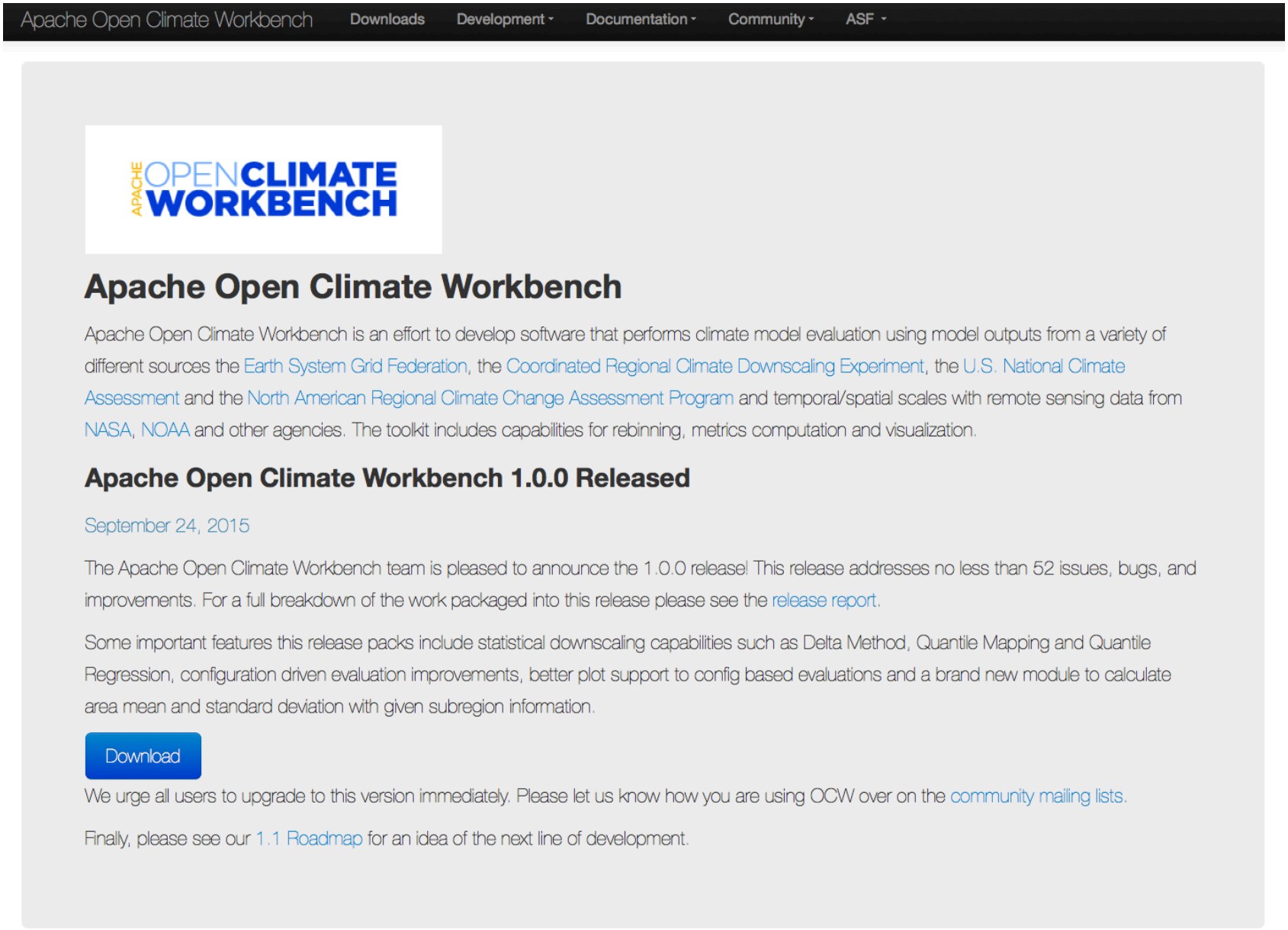


RCMES captures the entire workflow.

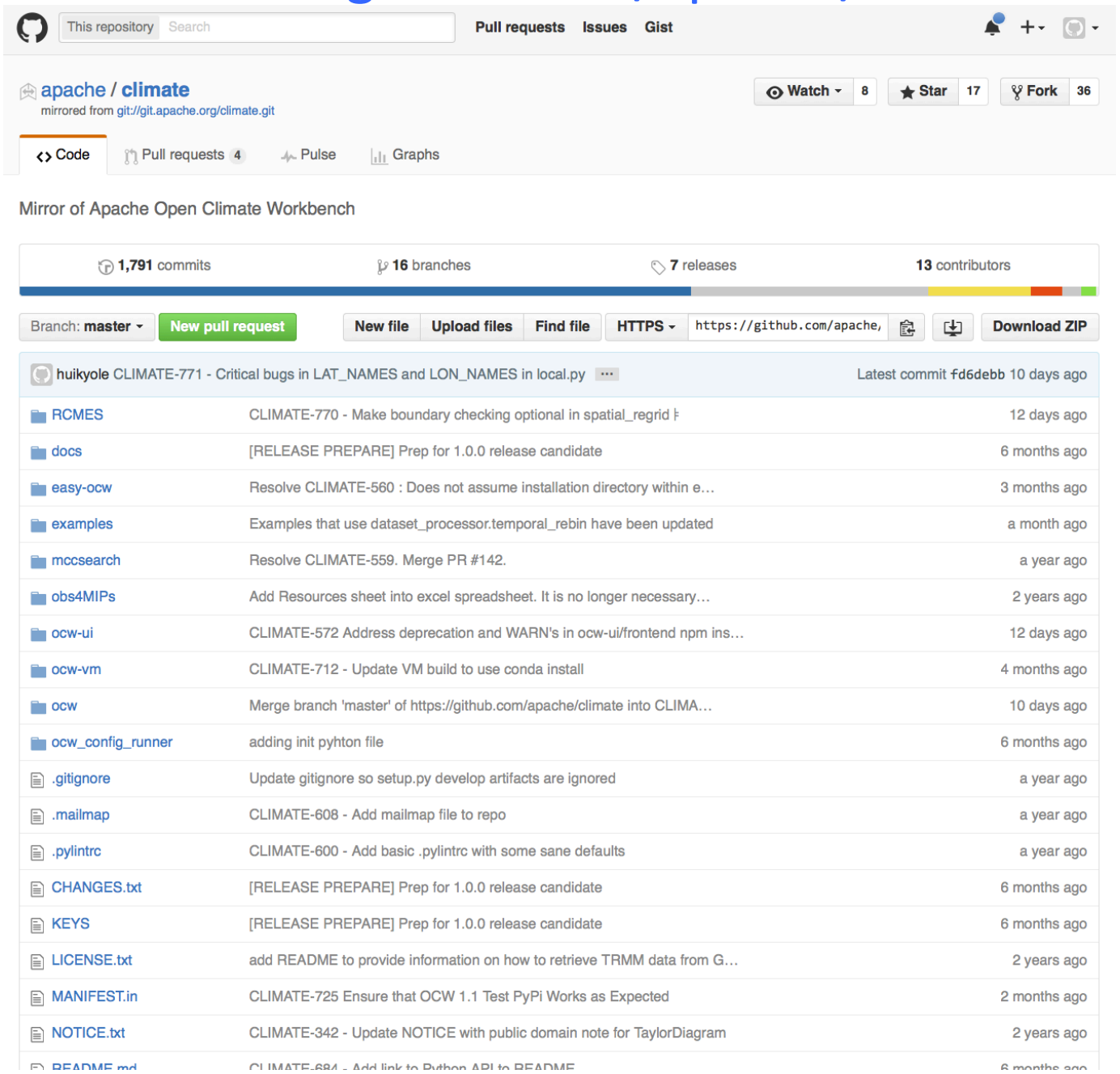
Another user can reproduce the same results using the captured workflow.

Apache Open Climate Workbench (OCW)

https://climate.apache.org/



Source at github.com/apache/climate



Running RCMES using configuration files:

a complete start-to-finish workflow to evaluate multi-scale climate models using observational data

load observation data (local, ESGF, RCMED)

dataset object1

dataset object 2

load model data (local, ESGF)

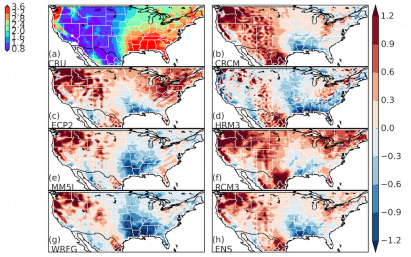
subset data (temporally & spatially)

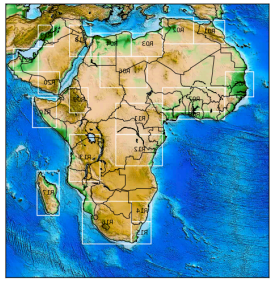
process data (filtering, averaging, regridding)

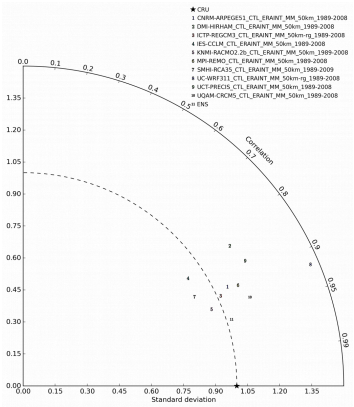
A netcdf file of the processed datasets

calculate metrics (bias, RMSE,…)

visualize results







Open Climate Workbench Libraries

data loader

dataset processor

metrics

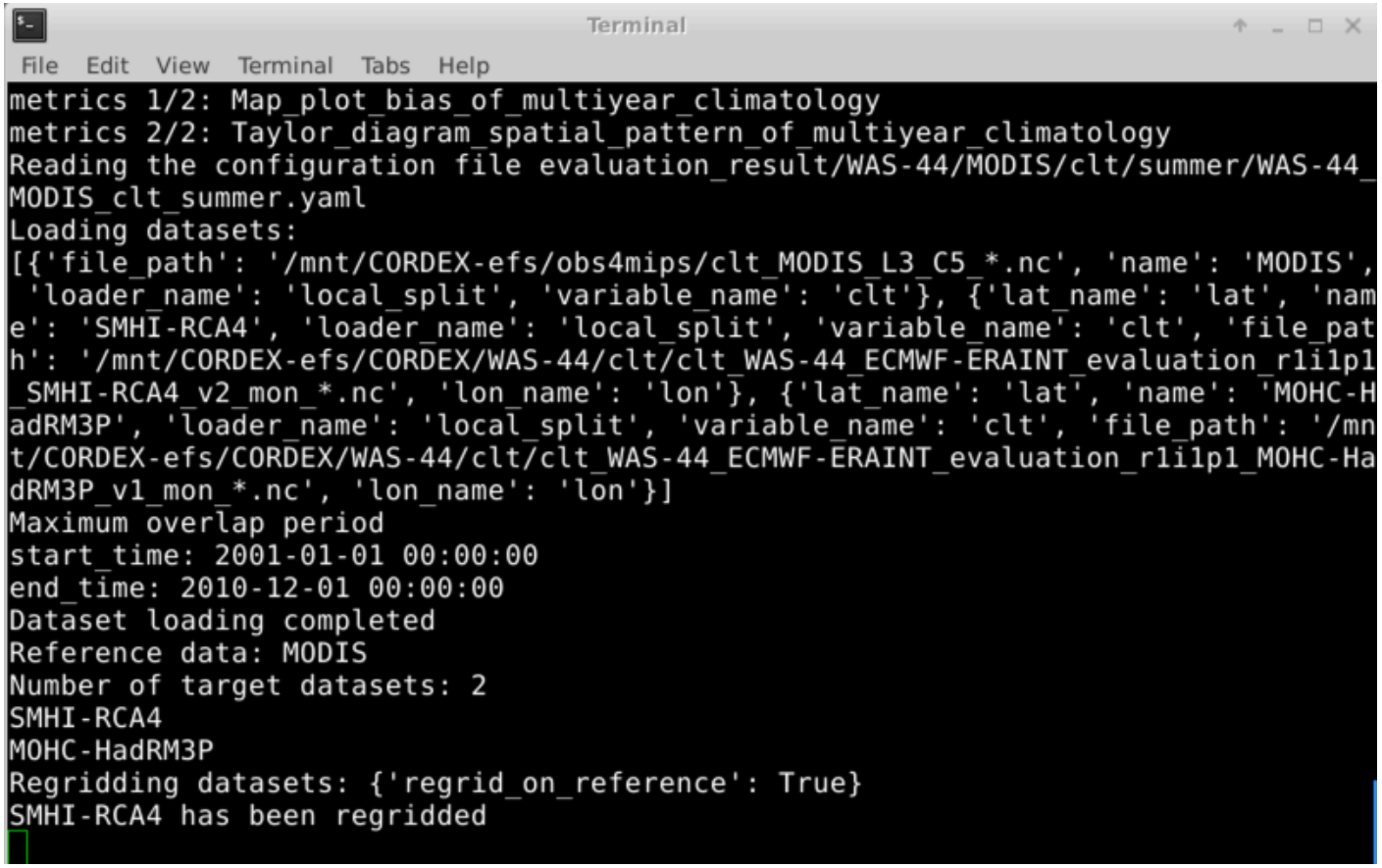
plotter

Running the systematic evaluation of CORDEX WAS simulations

1. Open Terminal and type `cd RCMES`

   `cd CORDEX_evaluation`

2. Type `./evaluate_WAS-44`

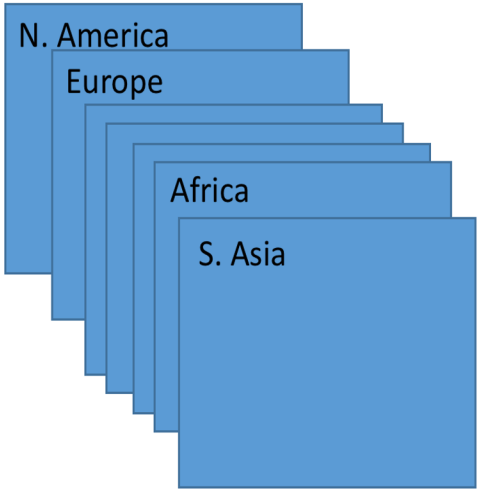


ERA-Interim Reanalysis Forced RCM simulations Available on ESGF

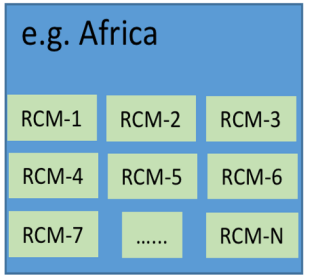
Schematic of Multi-Domain, Multi-Model and Multi-Variate CORDEX Model Evaluation with Obs4MIPs

Schematic of Multi-Domain, Multi-Model and Multi-Variate CORDEX Model Evaluation with Obs4MIPs

COREX has 14 Domains

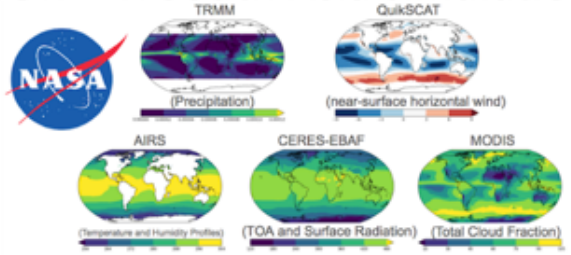


Each Domain has N RCMS



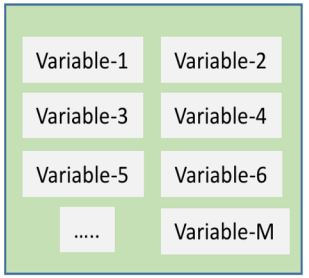
Mount an EFS storage to an AWS EC2 instance

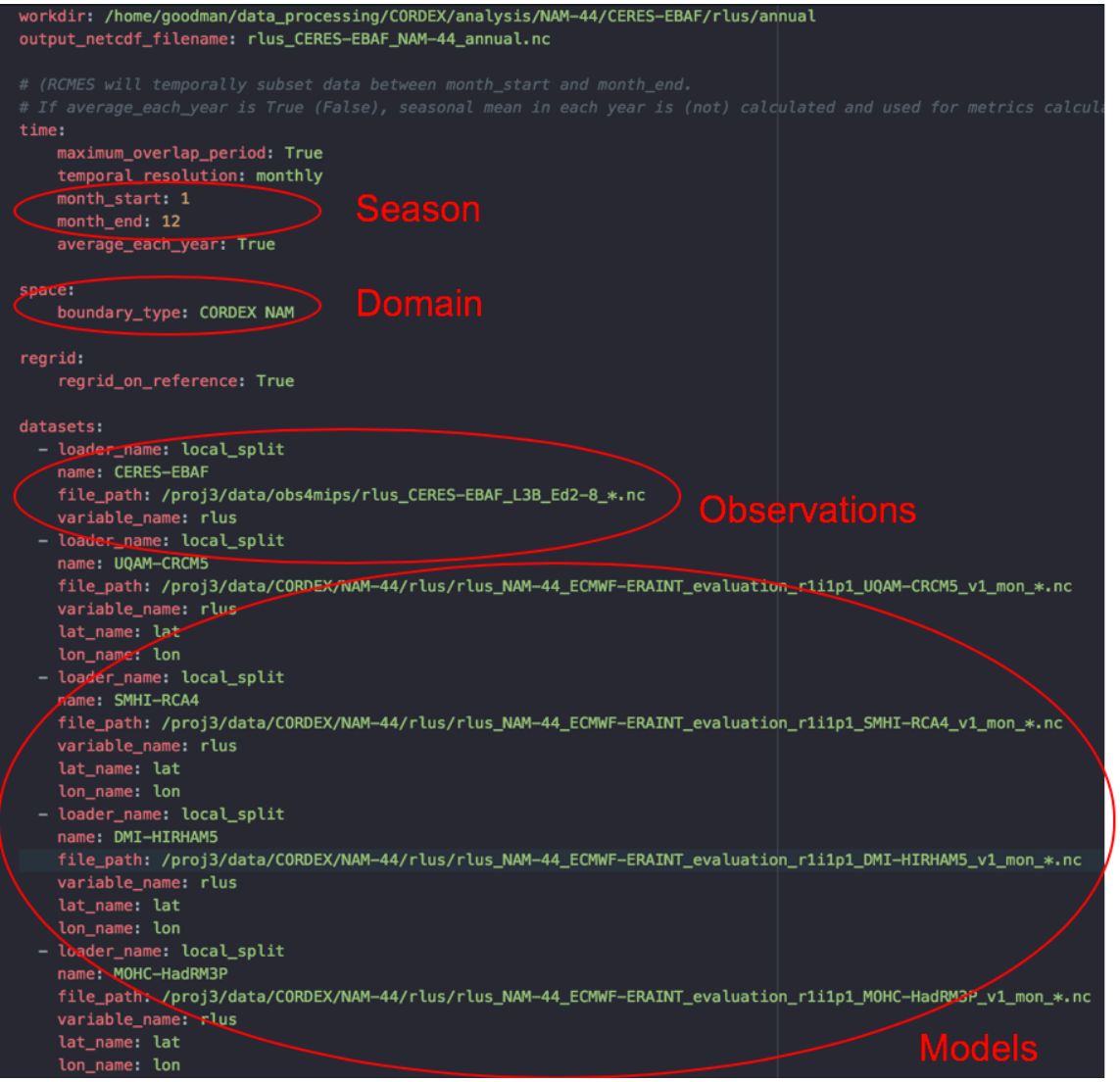
Elastic File System (EFS) storage containing obs4MIPs and CORDEX South Asia simulations



Model Evaluation

Each RCM has M Variables





Why we need "Systematic Evaluation"

- This Config file (namelist file) is necessary to run each evaluation combination (CORDEX Domain, Season and Variable), forming a large "evaluation matrix".
- 14 variables x 13 domains x 3 seasons x ~10 models > 5000 evaluations

- Writing that many Config files manually would be cumbersome/prohibitive.

Solution: Extract metadata from input filenames

rlus_NAM-44_ECMWF-ERAINT_evaluation_r1i1p1_UQAM-CRCM5_v1_mon_*.nc

Variable Domain **Model**

We can group all models and obs datasets together by common attributes (domain and variable) to form a unique evaluation, and therefore **automatically generate Config Files using only the dataset locations as user input**.

```
./evaluate_WAS-44
```

output folder name

```
python cordex.py WAS-44 evaluation_result /mnt/CORDEX-efs/obs4mips /mnt/CORDEX-efs/CORDEX/WAS-44
```

- Activity #1

: Correct biases in CORDEX RCM simulations

- Activity #2

: Evaluate CORDEX RCM simulations

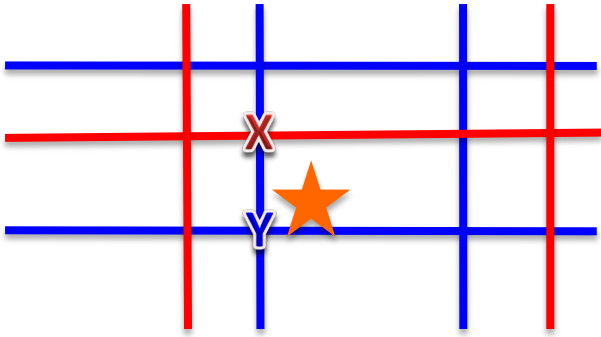
- Activity #3

: Pointwise Statistical downscaling using RCMES

- Activity #4

: Download and visualize the NEX-GDDP data

Statistical downscaling using RCMES



- To statistically downscale CMIP5 variables at a specific location (star marker), RCMES uses statistical relationship between the nearest model grid point data (X) and observation grid point data (Y)

  : **simultaneous correction of both bias and collocation**

$$Y = f(X)$$

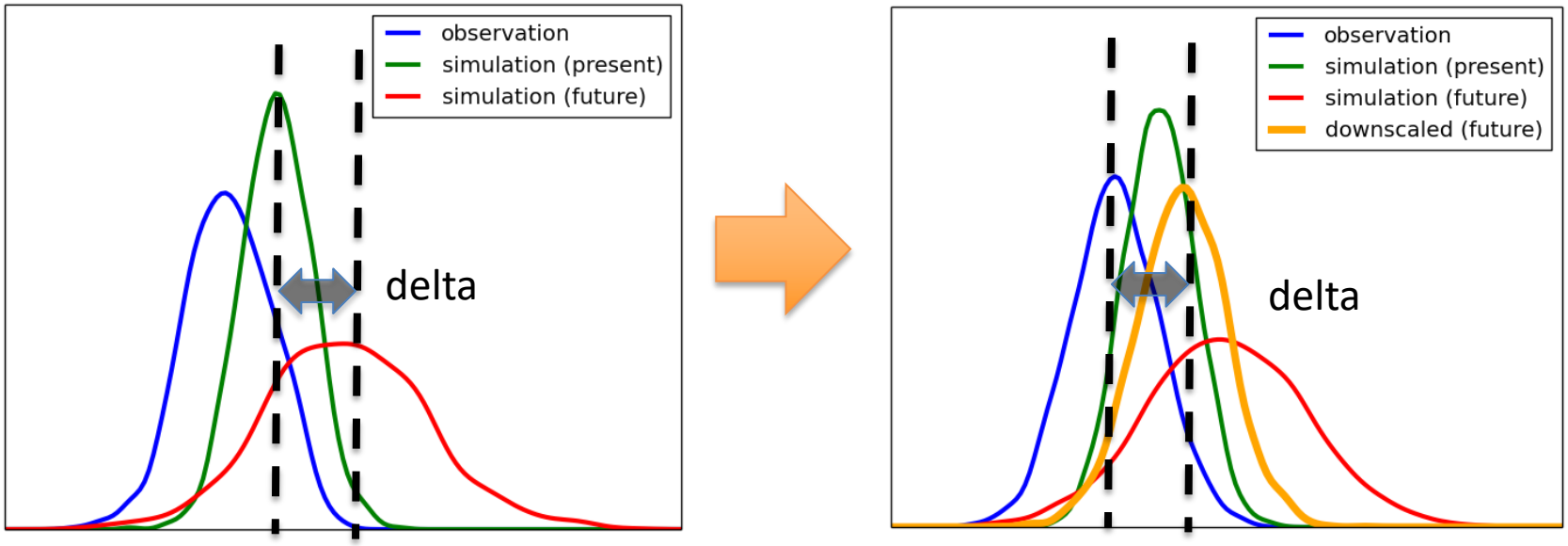
- Four different methods for model calibration (Stoner et al., 2013)
  - Delta method (addition)
  - Delta method (bias correction)
  - Quantile mapping
  - Asynchronous linear regression
- The observational datasets in RCMES database can be used to determine the observation-model relationship.

Delta method
(Delta addition)

Y_0: present observation, X_0: present simulation, X_1: future simulation

$$Y_1 = Y_0 + \bar{X}_1 - \bar{X}_0$$

- (future climate) = (present observation) + (mean difference between X_0 and X_1)

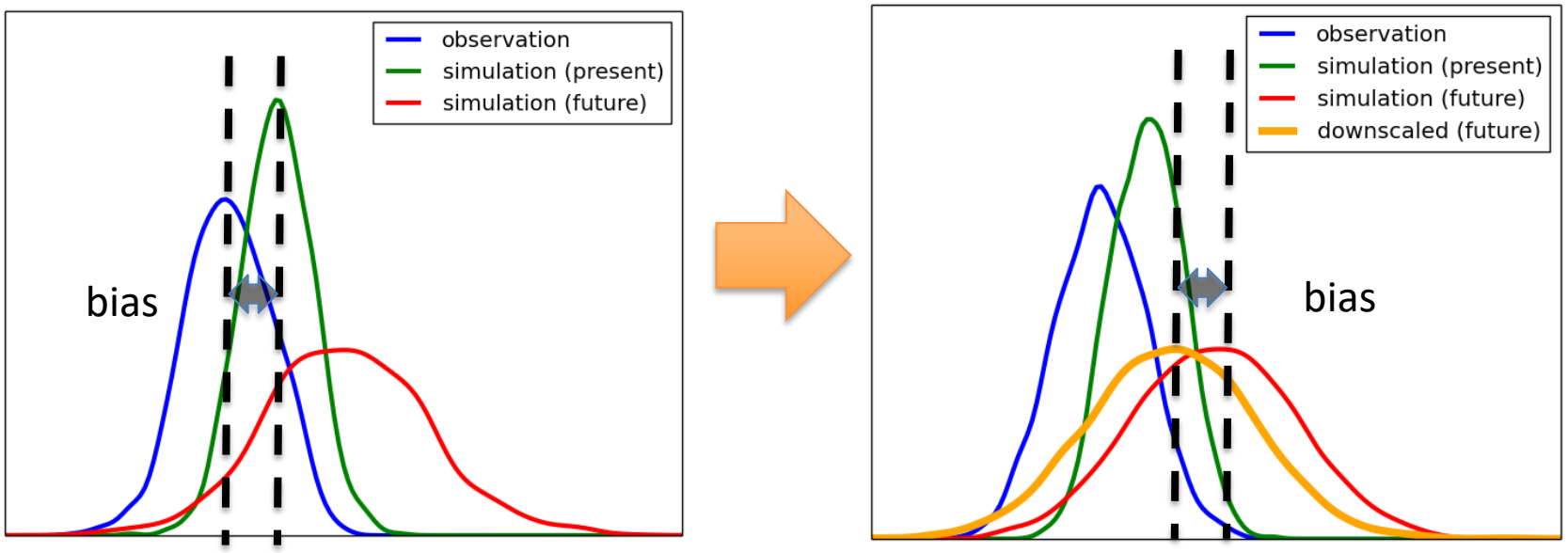


Delta method
(Bias correction)

Y_0: present observation, X_0: present simulation, X_1: future simulation

$$Y_1 = X_1 + \bar{Y_0} - \bar{X_0}$$

- (future climate) = (future simulation) + (mean bias)

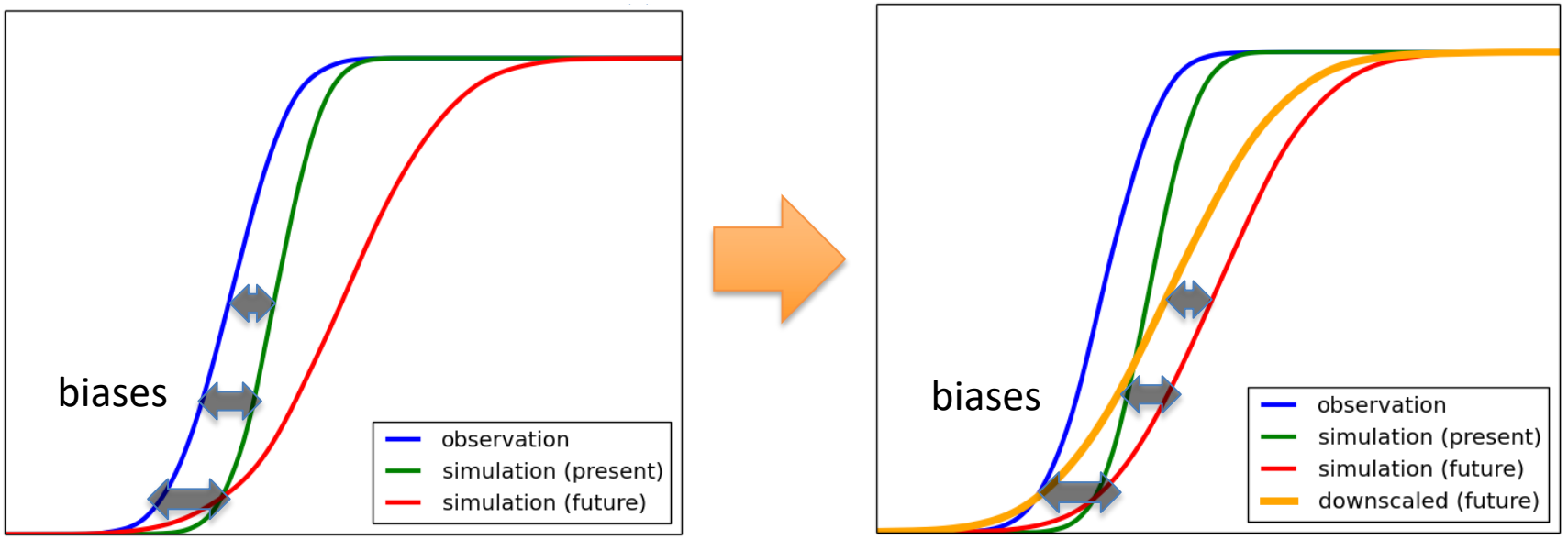


Quantile mapping

Y_0: present observation, X_0: present simulation, X_1: future simulation

$Y_1 = f(X_1)$ where f is bias correction function for each quantile ($Y_0 = f(X_0)$).

- (future climate) = (bias corrected future simulation)
- Bias is corrected for each quantile.

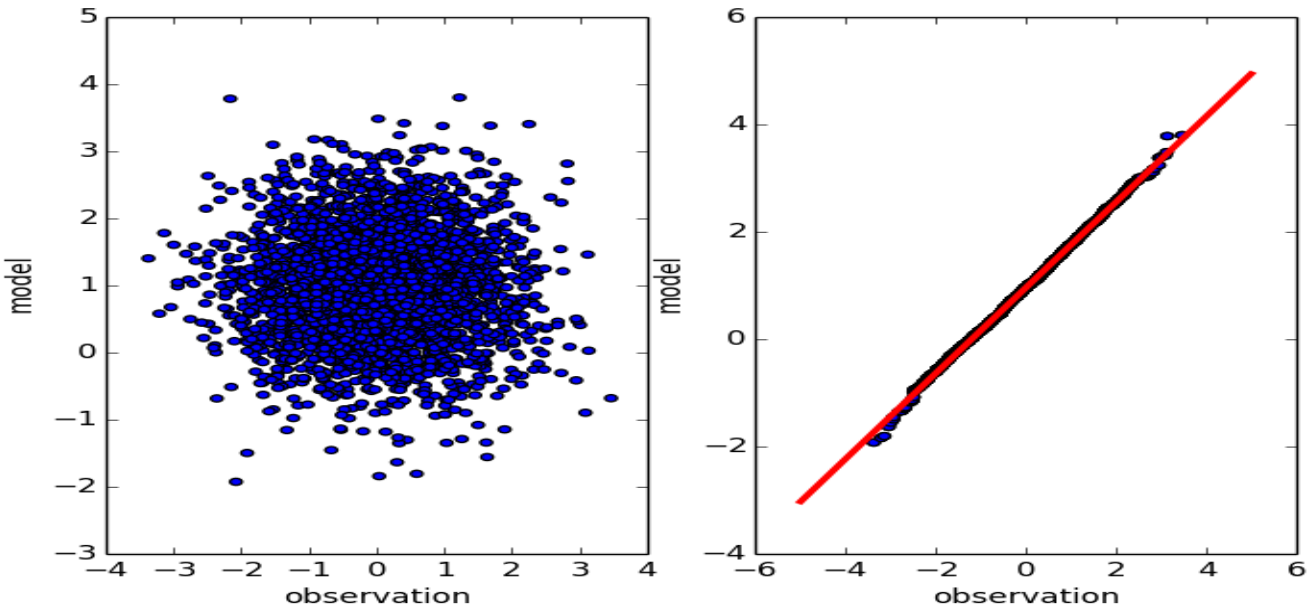


Asynchronous linear regression

Y_0: present observation, X_0: present simulation, X_1: future simulation
Y_0', X_0', X_1': sorted in ascending order

$Y_1' = a\dot{X}_1' + b$ where $Y_0' = a\dot{X}_0' + b$. a and b are the slope and intercept for the least square regression line.

- The linear relationship between observation and present simulation is determined after sorting them in ascending order.



Statistical Downscaling using RCMES

1. Open Terminal and `cd RCMES/statistical_downscaling/`
2. To run the statistical downscaling script, type

```
python run_statistical_downscaling.py MPI_tasmin_DJF.yaml
```

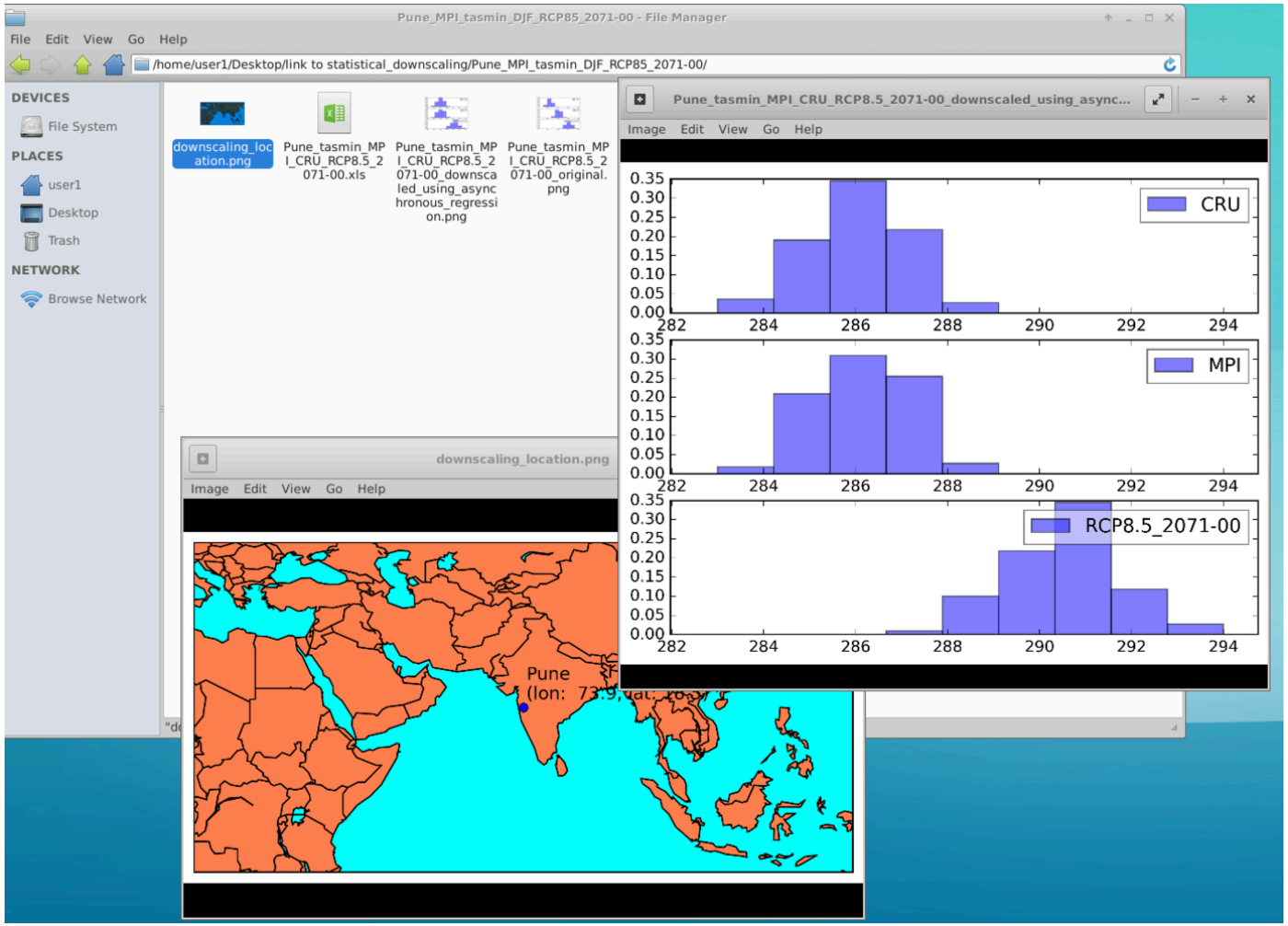
Python script

Configuration file

```
{13} [/home/ubuntu] % cd RCMES/statistical_downscaling/
data/  LaPaz_MPI_tas_JJA_RCP85_2071-00/  MIROC5_tasmax_JJA.yaml  MPI_tas_JJA.yaml  run_statistical_downscaling.py
{14} [/home/ubuntu/RCMES/statistical_downscaling] % python run_statistical_downscaling.py MPI_tas_JJA.yaml
Reading the configuration file  MPI_tas_JJA.yaml
Processing CRU  data
Loading ./data/tas_cru_monthly_1981-2010.nc into an OCW Dataset Object
CRU values shape: (times, lats, lons) - (360, 360, 720)

Loading ./data/tas_Amon_MPI_decadal1980_198101-201012.nc into an OCW Dataset Object
MPI values shape: (times, lats, lons) - (360, 96, 192)

RCP8.5_2071-00:MPI values shape: (times, lats, lons) - (360, 96, 192)

Temporal subsetting for the selected month(s)
Spatial aggregation of observational data near latitude 10.75 and longitude 106.67
Creating a statistical downscaling object
asynchronous_regression: Downscaling model output
Plotting results
Generating spreadsheet
{15} [/home/ubuntu/RCMES/statistical_downscaling] %
```

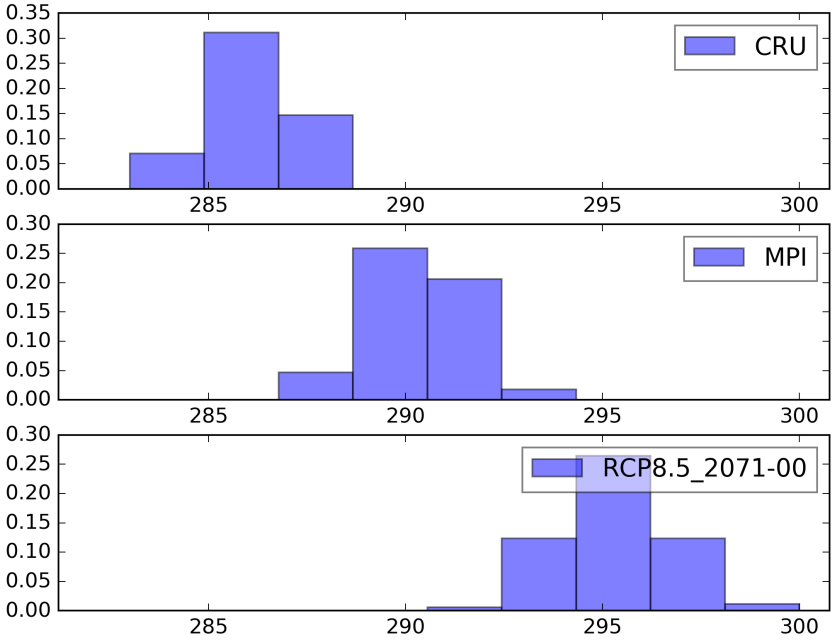
View the statistically downscaled tasmin results

- The results can be found in statistical_downscaling/Pune_tasmin_MPI_CRU_RCP8.5_2071-00 folder

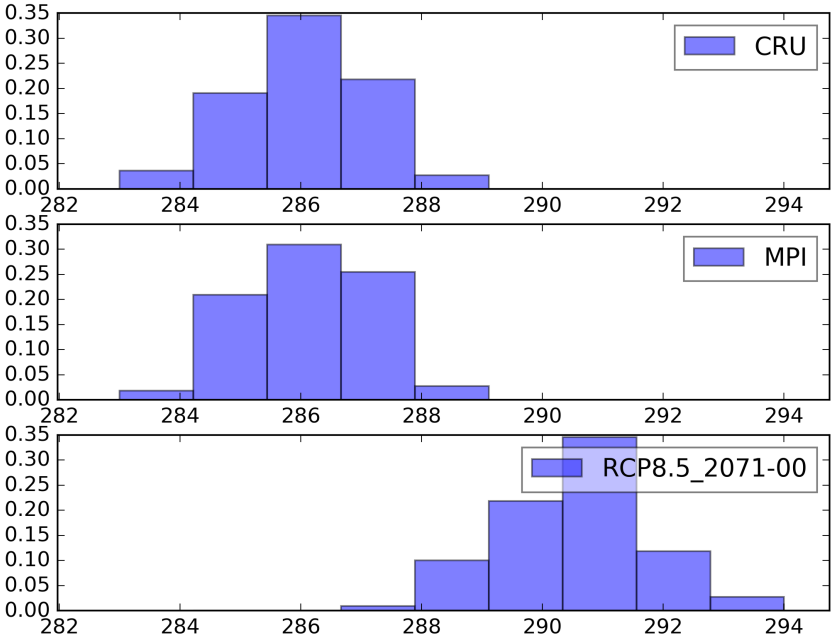


Quantile mapping of the daily minimum temperature for Pune in DJF

Original model output



Statistically downscaled model output



Run another example: taxmax in Pune

```
python run_statistical_downscaling.py MIROC5_taxmax_DEC.yaml
```

Python script

Configuration file

Make your own example by editing the yaml file

```
case_name: Pune_MIROC5_tasmax_DEC_RCP85_2071-00
downscaling_option: 3
location:
   name: Pune
   grid_lat: 18.5204
   grid_lon: 73.8567
 month_index: !!python/tuple [12]
reference:
   data_source: local
   data_name: CRUs
   path: ./data/tasmax_cru_monthly_1981-2010.nc
   variable: tasmax
model:
   data_name: MIROC5
   variable: tasmax
   present:
      path: ./data/tasmax_Amon_MIROC5_decadal1980_198
   future:
      scenario_name: RCP8.5_2071-00
      path: ./data/tasmax_Amon_MIROC5_rcp85_207101-210012.nc
```

Output folder name

Search Google with the keyword ‘latitude and longitude of XXX’

Season: December only in this case

(Options)

1. **IPSL**, **MPI**, and **MIROC5**
2. **tas**, **tasmin**, and **tasmax**
3. **RCP 4.5** and **8.5**
4. (**2041-2070**) and (**2071-2100**)

- Activity #1

: Correct biases in CORDEX RCM simulations

- Activity #2

: Pointwise Statistical downscaling using RCMES

- **Activity #3**

: Download and visualize the NEX-GDDP data

- Activity #4

: Analyze the bias corrected RCM output

NASA's Earth Exchange (NEX, https://nex.nasa.gov)

- NEX is a platform for scientific collaboration, knowledge sharing and research for the Earth science community.
- The new project, Open NEX, is aimed at making a number of important datasets more accessible.

NASA Earth Exchange
Global Daily Downscaled Projections (NEX-GDDP)

CMIP5 historical and RCP 4.5/8.5 simulations (from 21 models, 1950-2100)

Global Meteorological Forcing Dataset (observation, 1950-2005)

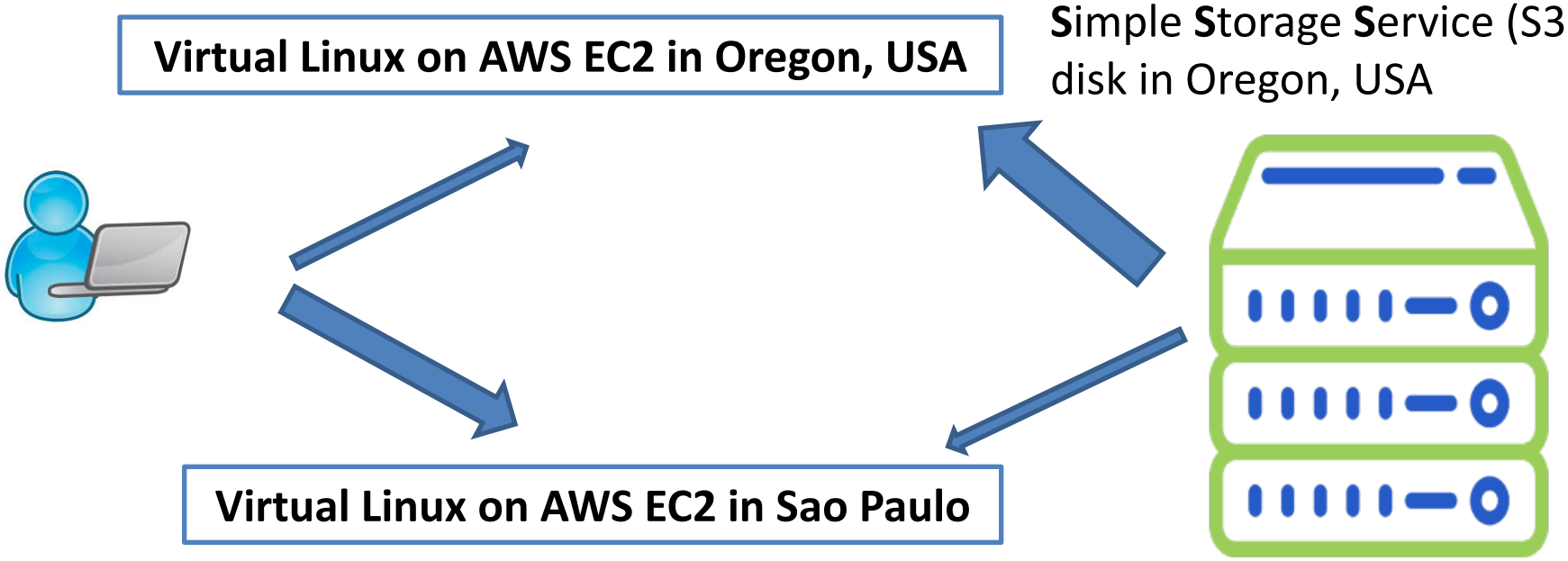
Bias-Correction Spatial Disaggregation (BCSD)

NEX-GDDP
: tasmax, tasmin, precipitation

Access to the statistically downscaled NEX-GDDP

256 Terabytes of the NEX-GDDP

Amazon **W**eb **S**ervice (AWS)
Simple **S**torage **S**ervice (S3)
disk in Oregon, USA



- The NEX S3 is mounted in your linux EC2.
- Open terminal and type `df -h`

What are inside s3://nasanex?

[/home/ubuntu] % `aws s3 ls s3://nasanex`

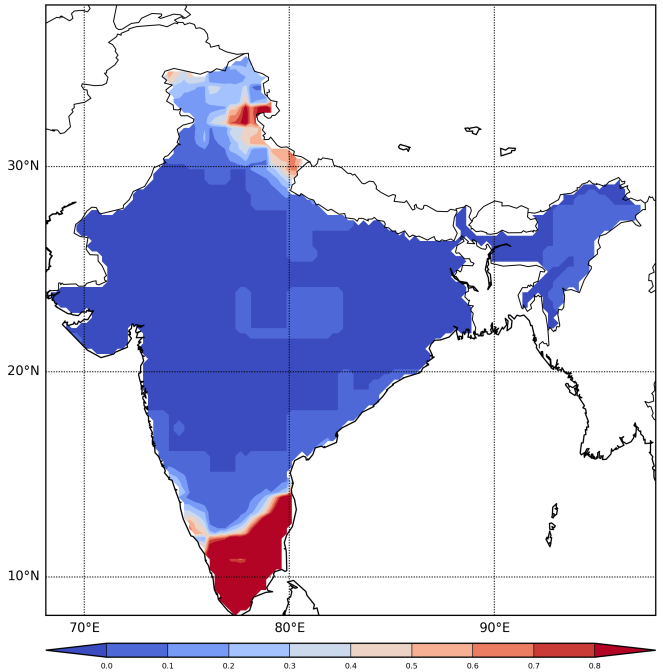
PRE AVHRR/
PRE CMIP5/
PRE LOCA/
PRE Landsat/
PRE MAIAC/
PRE MODIS/
PRE NAIP/
PRE NEX-DCP30/
PRE NEX-GDDP/

List, download and visualize NEX-GDDP

1. Open terminal and `cd NEX-GDDP`
2. `./list`
3. `./download` How fast!
4. `python plot_NEX-GDDP_example.py`

This script is an example of Open Climate Workbench, an open-source Python library that comprise RCMES.

Statistically downscaled precipitation from NorESM1-M model for December 2100



- Activity #1

: Correct biases in CORDEX RCM simulations

- Activity #2

: Pointwise Statistical downscaling using RCMES

- Activity #3

: Download and visualize the NEX-GDDP data

- **Activity #4**

: Analyze the bias corrected RCM output from Activity #1

- Activity #1

: Correct biases in CORDEX RCM simulations

- Activity #2

: Pointwise Statistical downscaling using RCMES

- Activity #3

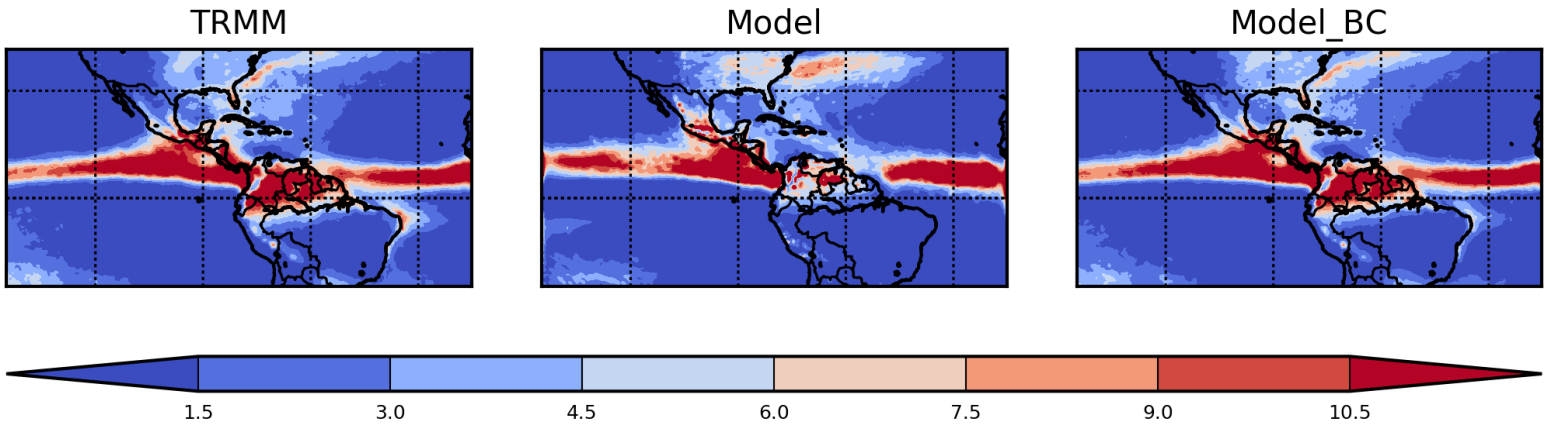
: Download and visualize the NEX-GDDP data

- **Activity #4**

: Analyze the bias corrected RCM output

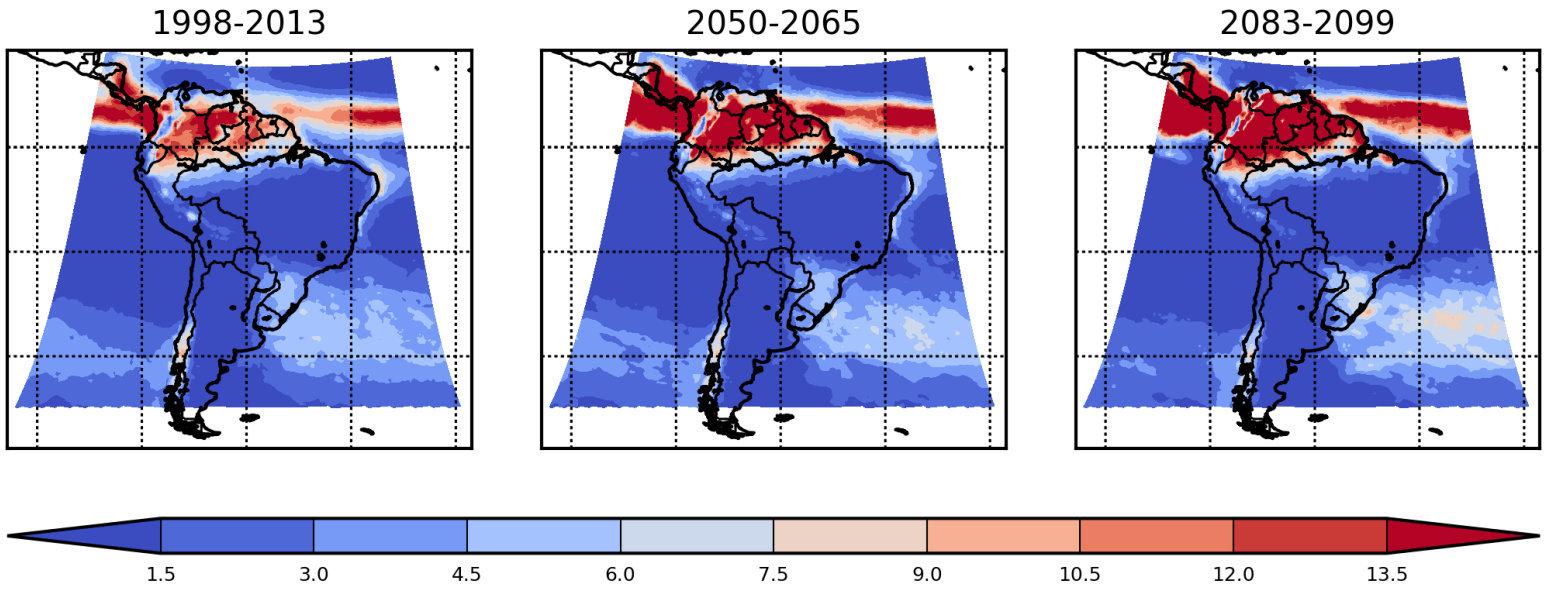
Compare TRMM, original simulation, and bias corrected simulation

1. Open terminal and `cd RCMES/analysis_examples`
2. `python check_bias_correction.py` ← OCW-based script



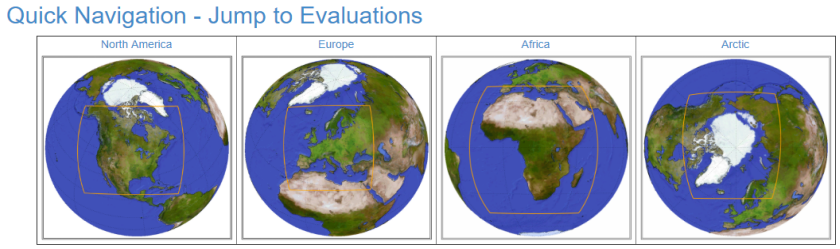
How does the bias-corrected precipitation look like in the future (in July-August, TRMM (1998-2013) vs. two bias-corrected simulations (2084-2099))?

```
python compare_present_and_future.py
```



Future Direction

- Development is ongoing…
  - Adding more metrics to assure traceability and reproducibility of model evaluation results.
  - Growing user and developer base by utilizing AWS and OpenNEX datasets.
- Develop a comprehensive model evaluation system for the United States National Climate Assessment and CORDEX.



Results

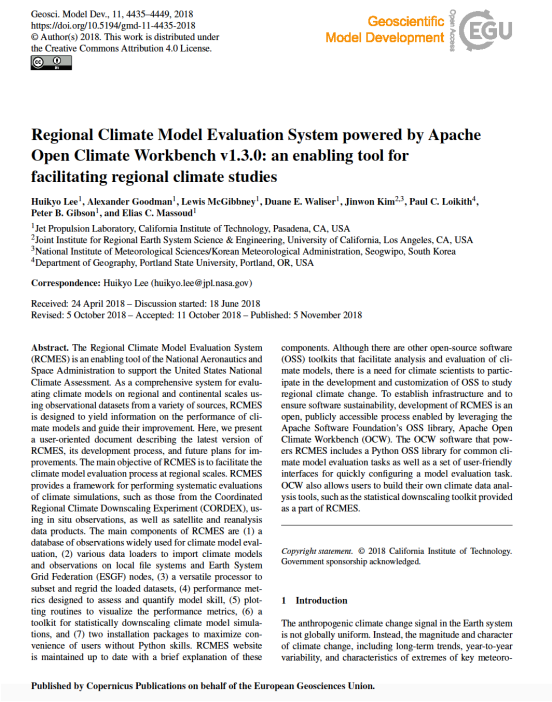
North America Evaluations

Reference Dataset	Variables	Results Page by Seasons		
CERES-EBAF	Downwelling Longwave Radiation (Surface)	Annual	Summer	Winter
	Upwelling Longwave Radiation (Surface)	Annual	Summer	Winter
	Upwelling Longwave Radiation (TOA)	Annual	Summer	Winter
	Downwelling Shortwave Radiation (Surface)	Annual	Summer	Winter
	Downwelling Shortwave Radiation (TOA)	Annual	Summer	Winter
	Upwelling Shortwave Radiation (Surface)	Annual	Summer	Winter
	Upwelling Shortwave Radiation (TOA)	Annual	Summer	Winter

https://rcmes.jpl.nasa.gov/content/cordex-evaluation

Where to find more information:

- http://rcmes.jpl.nasa.gov
- http://climate.apache.org/
- Email team members or dev@climate.apache.org
- https://nex.nasa.gov



Geosci. Model Dev., 11, 4435–4449, 2018
https://doi.org/10.5194/gmd-11-4435-2018

Geoscientific Model Development
EGU

Regional Climate Model Evaluation System powered by Apache Open Climate Workbench v1.3.0: an enabling tool for facilitating regional climate studies

Huikyo Lee[1], Alexander Goodman[1], Lewis McGibbney[1], Duane E. Waliser[1], Jinwon Kim[2,3], Paul C. Loikith[4], Peter B. Gibson[1], and Elias C. Massoud[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[2]Joint Institute for Regional Earth System Science & Engineering, University of California, Los Angeles, CA, USA
[3]National Institute of Meteorological Sciences/Korean Meteorological Administration, Seogwipo, South Korea
[4]Department of Geography, Portland State University, Portland, OR, USA

Correspondence: Huikyo Lee (huikyo.lee@jpl.nasa.gov)

Received: 24 April 2018 – Discussion started: 18 June 2018
Revised: 5 October 2018 – Accepted: 11 October 2018 – Published: 5 November 2018

Abstract. The Regional Climate Model Evaluation System (RCMES) is an enabling tool of the National Aeronautics and Space Administration to support the United States National Climate Assessment. As a comprehensive system for evaluating climate models on regional and continental scales using observational datasets from a variety of sources, RCMES is designed to yield information on the performance of climate models and guide their improvement. Here, we present a user-oriented document describing the latest version of RCMES, its development process, and future plans for improvements. The main objective of RCMES is to facilitate the climate model evaluation process at regional scales. RCMES provides a framework for performing systematic evaluations of climate simulations, such as those from the Coordinated Regional Climate Downscaling Experiment (CORDEX), using in situ observations, as well as satellite and reanalysis data products. The main components of RCMES are (1) a database of observations widely used for climate model evaluation, (2) various data loaders to import climate models and observations on local file systems and Earth System Grid Federation (ESGF) nodes, (3) a versatile processor to subset and regrid the loaded datasets, (4) performance metrics designed to assess and quantify model skill, (5) plotting routines to visualize the performance metrics, (6) a toolkit for statistically downscaling climate model simulations, and (7) two installation packages to maximize convenience of users without Python skills. RCMES website is maintained up to date with a brief explanation of these components. Although there are other open-source software (OSS) toolkits that facilitate analysis and evaluation of climate models, there is a need for climate scientists to participate in the development and customization of OSS to study regional climate change. To establish infrastructure and to ensure software sustainability, development of RCMES is an open, publicly accessible process enabled by leveraging the Apache Software Foundation's OSS library, Apache Open Climate Workbench (OCW). The OCW software that powers RCMES includes a Python OSS library for common climate model evaluation tasks as well as a set of user-friendly interfaces for quickly configuring a model evaluation task. OCW also allows users to build their own climate data analysis tools, such as the statistical downscaling toolkit provided as a part of RCMES.

1 Introduction

The anthropogenic climate change signal in the Earth system is not globally uniform. Instead, the magnitude and character of climate change, including long-term trends, year-to-year variability, and characteristics of extremes of key meteoro-

Published by Copernicus Publications on behalf of the European Geosciences Union.

Contact

Kyo Lee: huikyo.lee@jpl.nasa.gov